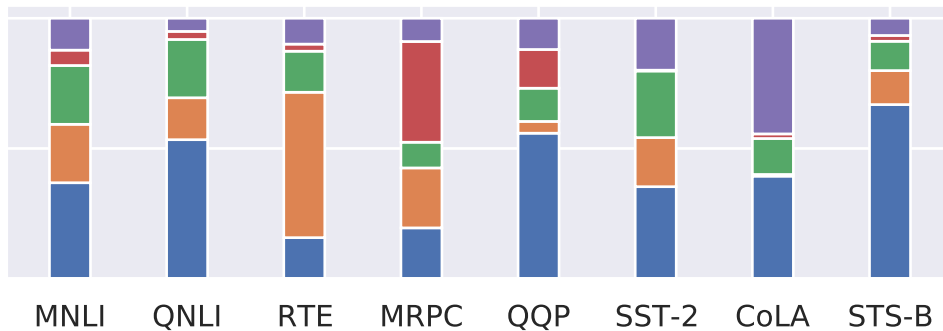Attention Patterns

Weight Normed Attention Patterns

Categories: Vertical + Diagonal, Vertical, Diagonal, Block, Heterogeneous

Datasets: MNLI, QNLI, RTE, MRPC, QQP, SST-2, CoLA, STS-B